Pervasive and Mobile Computing 6 (2010) 218–238

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

journal homepage: www.elsevier.com/locate/pmc

Optimising context data dissemination and storage in distributed pervasive computing systems

Ioanna Roussaki [a,*], Maria Strimpakou [a], Carsten Pils [b], Nikos Kalatzis [a], Nicolas Liampotis [a]

[a] National Technical University of Athens, School of Electrical and Computer Engineering, 9 Heroon Polytechneiou Street, 15773 Zographou, Athens, Greece

[b] Waterford Institute of Technology, Telecommunications Software & Systems Group, Waterford, Ireland

ARTICLE INFO

Article history:
Received 28 February 2008
Received in revised form 26 April 2009
Accepted 1 July 2009
Available online 19 July 2009

Keywords:
Context Distributed Database Management System
Context management & query languages
Mobile physical object handling
Context replication
Selective context data dissemination

ABSTRACT

Context management systems are expected to administrate large volumes of spatial and non-spatial information in geographical disperse domains. In particular, when these systems cover wide areas such as cities, countries or even the entire planet, the design of scalable storage, retrieval and propagation mechanisms is paramount. This paper elaborates on mechanisms that address advanced requirements, including support for distributed context databases management; efficient query handling; innovative management of mobile physical objects and optimization strategies for distributed context data dissemination. These mechanisms establish a robust spatially-enhanced distributed context management framework that has already been designed and carefully implemented and thoroughly evaluated.

1. Introduction

Network operators, sensor networks or even web resources capture valuable information such as device location and status, user profiles and movement patterns, network performance, etc., in order to provide enhanced telecommunication services to their clients. This data is considered to be vital context information [1] that can be exploited to customize services, to anticipate user intentions and to ultimately reduce human-to-machine interactions [2]. However, even though context information holds out the prospect of enhancing user experience and increasing revenues, disseminating it across distributed nodes is not straightforward. On the one hand, the various infrastructures that store and manage context data are heterogeneous, while there is no standardized interface that supports context information exchange. On the other hand, information, which cannot be retrieved when necessary, is valueless. In an open context marketplace, where a wide variety of information types is traded, context consumers are challenged by the discovery of the required context data. In this perspective, timely delivery of context information is crucial, due to the fact that most data sources provide real-time information. Thus, efficient approaches for distributed storage, retrieval and timely delivery of context data are essential for the success of context-aware computing systems. In particular, when context aware systems are not confined to a single building, but extend their functionality to cities, countries or even the entire planet, scalable and efficient data distribution mechanisms are paramount.

In order to support the provision of personalized and context aware services, irrespective of the user device and the underlying network, information originating from multiple context providers is usually required [3]. These context providers do not necessarily belong to the same legal entity or administrative domain. In such a federated context management

* Corresponding author. Tel.: +30 210 772 24 22; fax: +30 210 772 25 30.
E-mail addresses: nanario@telecom.ntua.gr (I. Roussaki), mstrim@telecom.ntua.gr (M. Strimpakou), cpils@tssg.org (C. Pils), nikosk@telecom.ntua.gr (N. Kalatzis), nliam@telecom.ntua.gr (N. Liampotis).

1574-1192/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.pmcj.2009.07.015

environment, user context information is a commodity that is protected by the administrative domains that possess it and therefore, stakeholders will hardly agree on a single Database Management System (DBMS). Though the stakeholders do not share context information, users have access to such information in case they have subscribed for it. This is, however, not a hard constraint. On the one hand, most DBMSs are relational databases that implement the SQL standard. On the other hand, standardizing the basic database structure is inevitable for the support of universal and efficient information exchange. Thus, a single virtual database that extends across all CM nodes must be in place. In this perspective, a Context Distributed Database Management System[1] (CDDBMS) should be established that would be responsible for processing context information as if this is maintained by a single database, while in fact the context information is stored and controlled by multiple administrative domains. Thus, the CDDBMS should provide the same functionality as a centralized database management system, making transparent to the context consumer that distributed query handling and interaction with remote context nodes are required.

The heterogeneity of context information imposes heavy requirements on the CDDBMS supporting pervasive systems. Undoubtedly, the most important and, at the same time, most critical context information that needs special treatment is location information. As users move across the globe, context-aware systems need to establish scalable mechanisms for ubiquitously supporting users irrespectively of their mobility pattern and current location. The introduced CDDBMS system should support effective mechanisms for dealing with mobile physical objects with regards to handling frequent updates of location information and is inspired by the mobile IP protocol. This paper initially focuses on presenting the details and advantages of the implemented CDDBMS.

Even though location is very critical in context-aware systems, there is a very wide variety of context information that is also necessary. In a pervasive world with millions of moving users and billions of interacting devices, an immense amount of heterogeneous context data is likely to be requested for delivery to remote nodes every second. In such an environment, apart from a quite small part of context data that will be strictly available to its owners (private), most context information will be a commodity to be traded among either a restricted consumers' group (semi-private) or frequently, among a large consumers' group (public). Even a well designed, distributed and scalable context management infrastructure would eventually fail in answering such continuously expanding demands for context information across remote context nodes. Especially, in cases where the requested context information constitutes a frequently updated piece of context data (e.g. room temperature) that should be conveyed to the client each time a new context value is available, the context management system, as well as the underlying network will be eventually flooded with the continuous context data updates to remote consumers. Each context consumer will have different requirements with regards to the context updates he/she is will willing to receive, e.g. he/she might desire to be informed once every day for the room temperature and not every second, thus making things even more complicated. Obviously, it is insufficient or even impossible to address each user's needs in full, anytime and anywhere. But if a group of users is interested in the same dynamic context information, sophisticated context dissemination mechanisms should be employed that will aim to simultaneously satisfy the entire group, although taking into consideration the individual characteristics and requirements of each user. The design and evaluation of such advanced context dissemination mechanisms is the main focus of this paper.

The rest of the paper is structured as follows. Section 2 elaborates on the Context Distributed Database Management System, briefly introducing the functional architecture of the established distributed context databases including the database schema, the supported query languages, as well as details about the mobile physical objects management. Section 3 presents an optimized context data dissemination mechanism caters for the cost-minimizing selective update of context replicas distributed in a multitude of context nodes throughout the network. More specifically, it focuses on establishing the general problem description, while also providing the formal problem statement and a brief review of the most relevant research work. Section 4 proposes heuristics algorithms for solving the context data dissemination problem distinguishing between four cases that are thoroughly described. Section 5 analyses in detail the conducted experiments and presents the obtained evaluation results. Section 6 performs a comparison of the proposed approach with the relevant state of the art work. Finally, in Section 7 conclusions are drawn, while an outline of the current status and future plans is provided.

2. The Context Distributed Database Management System

The CDDBMS is a peer-to-peer database comprising of node servers, each of which stores information about at least one predefined information domain. Such domains include for instance: user profiles, accounting information of a telecommunication operator, service advertisements or information related to a specific geographical area. Context producers store their data with respect to domains and consumers query information accordingly. Two domains are distinguished: Logical Domains (Logic-D) and Geographic Domains (Geo-D). Logic-Ds contain non-spatial information such as user profiles and are rather independent from other domains. Geo-Ds, however, carry strong spatial inclusion relationships. For example, the suburb is a geographic specialization of the city domain and thus, there is an inclusion relation between the City-Node and the Suburb-Node. Inclusion relations span a directed tree on the Geo-D space, where

[1] A DDBMS is a software system that permits the management of the distributed database and makes the distribution transparent to the users. A distributed database is a collection of multiple, logically interrelated databases distributed over a computer network [4]. Distributed database management systems are typically implemented as client-server systems or peer-to-peer systems [5].